Expected Shortfall
Expected shortfall is the expected value of Profit and Loss given the Loss is beyond VaR.

$$ES(X) = E(X \mid x \leq -VaR(X))$$

Expected Shortfall is the average of the left tail of the distribution, where the tail is defined as values whose CDF is less than or equal to Alpha.

Expected Shortfall is sometimes called:

1. Conditional VaR or CVaR
2. Average VaR or AVaR
3. Expected Tail Loss (ETL)

CVaR is the most widely used alternative, but can sometimes mean different things to different people. It is always good to ask for a definition when someone uses it.

Formally, we can construct this from the PDF and VaR

$$ES(X) = -\frac{1}{\alpha}\int_{-\infty}^{-VaR(X)} x\, f(x)\, dx$$

Numerically, given a sample of observations this is just the mean of all values less than or equal to VaR.

ES has closed form solutions for many distributions. Because most of what we do is simulation based, we will calculate ES numerically. In the case of the Delta Normal VaR, it is useful to have the equation for the normal.

$$ES_{\alpha}(X) = -\mu + \sigma\frac{f\left(\Phi^{-1}(\alpha)\right)}{\alpha}$$

Where $f(x)$ is the standard normal PDF, and $\Phi^{-1}(\alpha)$ is the standard normal quantile function.

Recognize that the expected shortfall for the normal is a linear function of the standard deviation. Just like VaR.

This doesn't hold for other distributions but Expected Shortfall has nice properties:

1. It is a coherent risk measure. That means ES is subadditive, like standard deviation, but unlike VaR
2. Expected Shortfall forms a convex surface when VaR does not. This means that optimizing based on ES is easier.

Expected Shortfall has gained in popularity. It gives a loss expectation instead of a minimum like VaR. It is more sensitive to the tails of the distribution than VaR. Fatter tails (more kurtosis) generally mean larger ES, when this is not always the case for VaR.

Model Based Simulation

Often we want to simulate from fitted models instead of just a distribution.

$$y = f(X, \epsilon)$$

To do this, we have to make distributional assumptions on X and ϵ. Easiest assumption is multivariate normal:

$$[X\ \epsilon] \sim N(\mu, \Sigma)$$

If $f(X, \epsilon)$ is OLS, then X and ϵ are independent and $E(\epsilon) = 0$.

$$y = X\beta + \epsilon$$

$$\begin{bmatrix} 1 & \rho_{12} & 0 \\ \rho_{21} & 1 & 0 \\ 0 & 0 & \end{bmatrix}$$: 2 Blocks

$$X \sim N\left(\mu_X, \Sigma_X\right),\ \epsilon \sim N\left(0, \sigma^2\right)$$

$$[X\ \epsilon] \sim N(\mu, \Sigma),\ \mu = \left[\mu_X\ 0\right],\ \Sigma = \underline{blockDiagonal}\left(\Sigma_x, \sigma^2\right)$$

always have block(s) independently from others, identical matrix is blockdiagonal matrix

We can simulate $[X, \epsilon]$, applying what we have learned previously.

In the case of OLS, this model fitting is not needed.

- OLS is a linear function of random normals
- A linear function of random normals is also a random normal
- Problem 1 on the first homework – OLS is the same things as just using a joint, multivariate, distribution on Y and X

What if we have n independent variables.

$$y_i = f(X, \epsilon_i)\ \forall\ i \in [1, n]$$

Let $\mathcal{E}$ be the matrix collection of ϵ_i vectors:

$$\mathcal{E} = \left[\epsilon_1, \epsilon_2, \dots, \epsilon_n\right]$$

Then Ɛ is

$$\mathcal{E} \sim N\left(0, \Sigma_{\epsilon}\right)$$

If all X variables are used in each model then:

$$[X\ \mathcal{E}] \sim N(\mu, \Sigma),\ \mu = \left[\mu_X\ 0\right],\ \Sigma = blockDiagonal\left(\Sigma_x, \Sigma_{\epsilon}\right)$$

If not all X variables are used in each model, then you have to calculate the full covariance matrix

$$[X\ \mathcal{E}] \sim N(\mu, \Sigma),\ \mu = \left[\mu_X\ 0\right],\ \Sigma = cov([X\ \mathcal{E}])$$

The diagonal elements will be the same, however, the off diagonal elements will not necessarily be 0. In practice, it is often easier to just calculate the covariance than to construct the block diagonal matrix.

The fact that everything is still normal holds even if there are multiple Dependent variables (Y) and multiple Independent Variables (X).

Recognize that a distributional assumption about a variable is a model

For example assuming $y \sim N\left(\mu, \sigma_{\epsilon}^2\right)$

$$y = f(X, \epsilon) = \mu + \sigma_{\epsilon} z$$

$$z \sim N(0, 1)$$

Breaking the Normality Assumption – Copulas
Often in finance, not all of the data are not normally distributed. Some might be close enough, but others are not. There are very few multivariate distributions with well defined properties. We need a way to combine disparate distributions into a single, coherent, simulation.

Copulas are multivariate cumulative distribution functions whose marginal probability distribution functions are uniform on [0,1]. They can be used to combine variables with different assumed distributions.

Sklar's Theorem forms the basis of copulas.

It states that every multivariate cumulative distribution function of random vector X

$$H(x_1, x_2, \dots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n)$$

Can be expressed in terms of its marginal distributions

$$F_i(x_i) = P(X_i \le x_i)$$

and some copula C

$$H(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n))$$

Recognize that a CDF $F_i(x_i) \in [0, 1]$.

A copula describes a joint distribution of variables, X, whose individual distributions are assumed. $F_i()$ does not have have to be the same distribution as $F_j()$.

We use the CDF to transform the variables into the range $[0, 1]$ and fit the copula.

Gaussian Copula

There are a number of different copulas. The most common is the Gaussian Copula. If you start to work with simulating financial data, looking into other copulas will be beneficial.

underestimate the risks

Given a Correlation matrix R, then the Gaussian copula is

$$C_R(X) = \Phi_R\left(\Phi^{-1}(F_1(x_1)), \Phi^{-1}(F_2(x_2)), \dots, \Phi^{-1}(F_n(x_n))\right)$$

$\Phi^{-1}(x)$ is the standard normal quantile function.
$\Phi_R(X)$ is the multivariate normal CDF given the correlation R.

These are all functions we have worked with and it makes it very simple to fit R

Fitting the copula:

1. For each variable $i \in 1 \ldots n$
    a. Transform the observation vector X_i in to a uniform vector U_i with $u_i \in [0, 1]$ using the CDF for x_i, $F_i\left(x_i\right)$.
    b. Transform the uniform vector U_i into a Standard Normal vector, Z_i using the normal quantile function.
2. Calculate the correlation matrix of Z

When calculating Z, it is generally better to use Spearman correlations. We are often dealing with distributions with large outliers and Spearman is a more robust correlation measure with outliers.

Spearman correlation for reference

$$r_s = \rho_{R(x),R(y)} = \frac{cov(R(x),R(y))}{\sigma_{R(x)}\sigma_{R(y)}} = \frac{\sum_{i=1}^{n}\left(R\left(x_i\right) - \overline{R(x)}\right)\left(R\left(y_i\right) - \overline{R(y)}\right)}{\sqrt{\sum_{i=1}^{n}\left(R\left(x_i\right) - \overline{R(x)}\right)^2}\sqrt{\sum_{i=1}^{n}\left(R\left(y_i\right) - \overline{R(y)}\right)^2}}$$

Where
$\rho_{x,y}$ is the Pearson coefficient of x and y
$R(x)$ is the rank function
$\overline{R(x)}$ is the average of the rank function

Because Spearman uses Rank correlations and the quantile function is monotonic, we can skip step 2.a (calculate Z values), and calculate Spearman correlations from the U matrix.

Simulating from the Copula
Once the copula is fit, we use the multivariate normal simulation techniques we learned previously. This generates correlated standard normals. We go backwards through the transform in the previous section to arrive at the simulated values of X.

1. Simulation $NSim$ draws from the multivariate normal
2. For each variable $Z_i \in 1\ldots n$.
    a. Transform Z_i into a uniform variable using the standard normal CDF, $\Phi\left(z_i\right) = U_i$.
    b. Transform U_i into the fitted distribution using the quantile of the fitted distribution, $F_i^{-1}\left(U_i\right) = X_i$

Putting it together:

Example Problem.

- Calculate VaR on a portfolio holding 1 share of each stock in DailyPrices.csv.
- Model prices as a linear function of SPY with T distributed Errors.
- Model SPY as normally distributed.
- Assume all stocks and SPY have 0 expected return.
- Calculate VaR for Each Stock as well as the total portfolio